Do not Enter
MW 11/21/02

means due, for example, to differences in calibration, then the input signals will be pre-normalized to the same nominal mean values during initial operation.

The specific goal of the A1 engine is to declare system 1 or system 2 degraded if the drift in Y is sufficiently large that the sequence of observations appears to be distributed about means + M or - M, where M is a preassigned system distribution magnitude. The SPRT provides a quantitative framework that enables us to decide between two hypotheses, H and H2, namely:

H1: Y is drawn from a Gaussian product distribution function (PDF) with means M and variance σ^2.

H2: Y is drawn from a Gaussian PDF with mean O and variance σ^2.

If it is supposed that H1 or H2 is true, we wish to decide for H1 or H2 with probability (1-β) or (1-α) respectively, where α and β represent the error (misidentification) probabilities.

From the theory described by Wald and Wolfowitz in "*Optimum Character of the Sequential Probability Ratio Test,*" *Ann. Math. Stat.*, 19,326 (1948), the most powerful test depends on the likelihood ratio 1_n, where

Probability of observed sequence given H1 true.

$$l_n = \frac{y_1, y_2, \ldots, y_n}{y_1, y_2, \ldots, y_n}$$

Probability of observed sequence given H2 true.

After n observations have been made, the sequential probability ratio is just the product of the probability ratio is just the product of the probability ratios for each step:

$$l_n = (PR1) \cdot (PR2) \cdot (PR3) \cdot \ldots (PR_n)$$

or

$$l_n = \prod_{i=1}^{n} \frac{f(yi|H_1)}{f(yi|H_2)},$$

where $F(y_i|H)$ is the distribution of the random variable y.

the Wald-Wolfowitz theory operates as follows:

Continue sampling as long as

$$A < 1n < B \qquad (1)$$

Stop sampling and decide $H1$ as soon as $1_n \geq B$, and stop sampling and decide $H2$ as soon as $1_n \leq A$. The acceptance thresholds are related to the error (misidentification) probabilities by the following expressions:

$$A = \frac{\beta}{1-\alpha} \text{ and } B = \frac{1-\beta}{\alpha} \qquad (2)$$

where

α = probability of accepting $H1$ when $H2$ is true (false alarm probability)

β = probability of accepting $H2$ when $H1$ is true (missed alarm probability)

Assuming the random variable y_k is normally distributed, the likelihood that $H1$ is true (mean M, variance σ^2) is given by

$$L(y1, y2, y3, \ldots yn|H1) = \frac{1}{(2\pi)^{n/2}\sigma^2} \exp\left[-\frac{1}{2\sigma^2}\left(\sum_{k=1}^{n} y_k{}^2 - 2\sum_{k=1}^{n} y_k M = \sum_{k=1}^{n} M^2\right)\right]. \quad (3)$$

Similarly for $H2$ (means o, variance σ^2),

$$L(y_1, y_2, y_3, \ldots y_n|H2) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^{n} y_k{}^2\right) \quad (4)$$

The ratio of equations (3) and (4) gives the likelihood ratio 1_n; where 1_n is expressed as

$$l_n = \exp\left[\frac{-1}{2\sigma^2}\sum_{k=1}^{n} M(m - 2y_k)\right] \quad (5)$$

combining equations 1, 2 and 5, and taking the natural logs, gives

$$l_n[\beta/(1-\alpha)] < \frac{-1}{2\sigma^2}\sum_{k=1}^{n} M(M - 2y_k) < l_n[(1-\beta)\alpha]$$

where

$$SPRT = \frac{-1}{2\sigma^2}\sum_{k-1}^{n} M(M - 2y_k) \text{ or}$$

then the sequential sampling and decision strategy can be concisely represented as

If SPRT$\leq 1_n(\beta/1-\alpha)$ accept H2

If $1_n[\beta/(1-\alpha)] < \text{SPRT} < 1_n[(1-\beta)/\alpha]$,

continue sampling

If SPRT$\geq 1_n[(1-\beta)/\alpha]$ accept H1.

The SPRT analysis formulated here cannot be applied directly to non-Gaussian signals. For applications to nuclear system signals contaminated by non-Gaussian noise, an attempt should first be made to pretreat the input signals with a normalizing transformation.

For applications where (a) one requires a high degree of assurance that a system is functioning within specifications and (b) there is not a large penalty associated with false alarms, it is not uncommon to specify a B (missed alarm probability) that is much smaller than A (false alarm probability). In safety critical systems one may be more willing to incur a false alarm than a missed alarm. For applications where a large cost penalty is incurred with any false alarms, it is desirable to keep both A and B small.

The trade-off that must be considered before one specifies arbitrarily small values for A and B is the effect this may have on the sensitivity and maximum decision time needed by the SPRT to annunciate a disturbance. The desired sensitivity of the SPRT is fixed by specification of M, the system disturbance magnitude. For a given value of M, the average sample number required to reach a decision is influenced by A and B and also by the variance associated with the signals being monitored. It takes longer to identify a subtle change in a process characterized by a low signal-to-noise ratio than in one with a high signal-to-noise ratio.

The non-stationary version of the SPRT algorithm is a slightly modified version of Wald's SPRT. In the non-stationary case, the failure magnitude, M, reference signal (or

mean), μ, and the reference variance, σ^2, are sample dependent. Therefore, the non-stationary SPRT equation becomes

$$SPRT(n) = SPRT(n-1) + \frac{M(n)}{\sigma^2(n)}\left((y(n) - \mu(n)) - \frac{M(n)}{2} \right),$$

where n = 1,2,...,L and L is the length of the length equalized signals. In this case, y(n) is the length of the equalized monitored signal, $\mu(n)$ is the corresponding reference signal generated during the training phase and $\sigma^2(n)$ is the variance of each point in $\mu(n)$.

The bounded angle ratio test (hereinafter BART) mentioned above is employed in systems with more than two variables, as shown in FIG. 9. For example, BART can be used on an actual sensor signal exhibiting non-white characteristics, such as for example, on sensor signals from the primary pump #2 of the EBR-II nuclear reactor at Argonne National Laboratory (West) in Idaho. In such a case, the signal can be a measure of the pump's speed over a given amount of time. In such a situation, one can use a nonlinear multivariate regression technique that employs an N Dimensional Space (known in vector calculus terminology as hyperspace) to model the relationships between all of the variables. This regression procedure results in a nonlinear synthesized estimate for each input observation vector based on the hyperspace regression model. The nonlinear multivariate regression technique is centered around the hyperspace BART operator that determines the element by element and vector to vector relationships of the variables and observation vectors, given a set of system data that is recorded during a time period when everything is functioning correctly.

In the BART method described in FIG. 9., the method is also split into a training phase and a monitoring phase. The first step in the training phase is to acquire a data

matrix continuing data samples from all of the sensors (or data sources) used for monitoring the system that are coincident in time and are representative of normal system operation. Then the BART parameters are calculated for each sensor (X_{med}, X_{max} and X_{min}). Here X_{med} is the median value of a sensor. The next step is to determine the similarity domain height for each sensor (h) using the BART parameters X_{med}, X_{max} and X_{min}. Once these parameters are calculated a subset of the data matrix is selected to create a model matrix (H) that is used in the BART estimation calculations. Here, H is an NxM matrix where N is the number of sensors being monitored and M is the number of observations stored from each sensor. The last steps taken during the training phase are the SPRT parameters calculations. The calculations are analogous to the calculations in the other methods, except that now the standard deviation value used to calculate SDI is obtained from BART estimation errors from each sensor (or data source) under normal operating conditions.

During the BART monitoring phase, a sample vector is acquired at each time step t, that contains a reading from all of the sensors (or data sources) being used. Then the similarity angle (SA) between the sample vector and each sample vector stored in H is calculated. Next an estimate of the input sample vector Y is calculated using the BART estimation equations. The difference between the estimate and the actual sensor values is then used as input to the SPRT module. Each difference is treated separately so that a decision can be made on each sensor independently. This method is described in more detail hereinafter.

In this preferred embodiment of FIG. 9 of the invention, the method measures similarity between scalar values. BART uses the angle formed by the two points under

comparison and a third reference point lying some distance perpendicular to the line formed by the two points under comparison. By using this geometric and trigonometric approach, BART is able to calculate the similarity of scalars with opposite signs.

In the most preferred form of BART an angle domain must be determined. The angle domain is a triangle whose tip is the reference point (R), and whose base is the similarity domain. The similarity domain consists of all scalars which can be compared with a valid measure of similarity returned. To introduce the similarity domain, two logical functional requirements can be established:

0 The similarity between the maximum and minimum values in the similarity domain is 0, and

1 the similarity between equal values is 1.

Thus the similarity range (i.e. all possible values for a measure of similarity), is in the range 0 to 16 inclusive.

BART also requires some prior knowledge of the numbers to be compared for determination of the reference point (R). Unlike a ratio comparison of similarity, BART does not allow "factoring out" in the values to be compared. For example, with the BART methodology the similarity between 1 and 2 is not necessarily equal to the similarity between 2 and 4. Thus, the location of R is vital for good relative similarities to be obtained. R lies over the similarity domain at some distance h, perpendicular to the domain. The location on the similarity domain at which R occurs (X_{med}) is related to the statistical distribution of the values to be compared. For most distributions, the median or mean is sufficient to generate good

results. In a preferred embodiment the median is used since the median provides a good measure of data density and is resistant to skewing caused by large ranges of data.

Once X_{med} has been determined, it is possible to calculate h. In calculating h, it is necessary to know the maximum and minimum values in the similarity domain. (X_{max} and X_{min} respectively) for normalization purposes the angle between X_{min} and X_{max} is defined to be 90°. The conditions and values defined so far are illustrated in FIG. 10. From this triangle it is possible to obtain a system of equations and solve for h as shown below:

$$\begin{aligned} &c\text{-}X_{med}\text{-}X_{min} \\ &d\text{-}X_{max}\text{-}X_{med} \\ &a^2\text{-}c^2=h^2 \\ &b^2=d^2+h^2 \\ &(c+d)^2=a^2+b^2 \\ &(c+d)^2=c^2+d^2+2h^2 \\ &h^2=cd \\ &h=\sqrt{cd} \end{aligned} \qquad (19)$$

Once h has been calculated the system is ready to compute similarities. Assume that two points: X_0 and X_1 ($X_0 \le X_1$) are given as depicted in FIG. 11 and the similarity between the two is to be measured. The first step in calculating similarity is normalizing X_0 and X_1 with respect to X_{med}. This is done by taking the euclidean distance between X_{med} and each of the points to be compared. Once X_0 and X_1 have been normalized, the angle $\angle X_0RX_1$ (hereinafter designated θ) is calculated by the formula:

$$\theta = \text{ArcTan}(X_1|h) = \text{ArcTan}(X_0|h) \qquad (20)$$

After θ has been found, it must be normalized so that a relative measure of similarity can be obtained that lies within the similarity range. To ensure compliance with functional requirements (A) and (B) made earlier in this section, the relative similarity angle (SA) is given by:

$$SA = 1 - \frac{\theta}{90^\circ} \qquad (21)$$

Formula (21) satisfies both functional requirements established at the beginning of the section. The angle between X_{min} and X_{max} was defined to be 90°, so the similarity between X_{min} and X_{max} is 0. Also, the angle between equal values is 0° The SA therefore will be confined to the interval between zero and one, as desired.

To measure similarity between two vectors using the BART methodology, the average of the element by element SAs are used. Given the vectors x_1 and x_2 the SA is found by first calculating S_i for i= 1,2,3...n for each pair of elements in x_1 and x_2 i.e.,

$$\text{if } \underline{x_1} = [X_{11} X_{12} X_{13} \cdots X_{1n}] \text{and } \underline{x_2} = [X_{21} X_{22} X_{23} \cdots X_{2n}]$$

The vector SA $\underline{\Gamma}$ is found by averaging over the Si's and is given by the following equation:

$$\Gamma = \frac{1}{n} \sum_{i=1}^{n} S_i \qquad (22)$$

In general, when given a set of multivariate observation data from a process (or other source of signals), linear regression could be used to develop a process model that relates

all of the variables in the process to one another. An assumption that must be made when using linear regression is that the cross-correlation information calculated from the process data is defined by a covariance matrix. When the cross-correlation between the process variables is nonlinear, or when the data are out of phase, the covariance matrix can give misleading results. The BART methodology is a nonlinear technique that measures similarity instead of the traditional cross-correlation between variables. One advantage of the BART method is that it is independent of the phase between process variables and does not require that relationships between variables be linear.

If there is a random observation vector $\underline{y}$ and a known set of process observation vectors from a process P, it can be determined if $\underline{y}$ is a realistic observation from a process P by combining BART with regression to form a nonlinear regression method that looks at vector SAs as opposed to euclidean distance. If the know observation vectors taken from P are given by

$$H = \left[\begin{pmatrix} h_{11} \\ h_{21} \\ h_{31} \\ \vdots \\ h_{k1} \end{pmatrix} \begin{pmatrix} h_{12} \\ h_{22} \\ h_{32} \\ \vdots \\ h_{k2} \end{pmatrix} \cdots \begin{pmatrix} h_{1m} \\ h_{2m} \\ h_{3m} \\ \vdots \\ h_{km} \end{pmatrix} \right] = \left[\underline{h}_1 \ \underline{h}_2 \cdots \underline{h}_m \right] \quad (23)$$

where H is k by m (k being the number of variables and m the number of observations), then the closest realistic observation vector to $\underline{y}$ in process P given H is given by

$$\underline{y} = H\underline{w} \quad (24)$$

Here $\underline{w}$ is a weighting vector that maps a linear combination of the observation vectors in H to the most similar representation of $\underline{y}$. The weighting vector $\underline{w}$ is calculated by

combining the standard least squares equation form with BART. Here θ stands for the SA operation used in BART.

$$\underline{w} = (H' \oplus H)^{-1} H' \oplus \underline{y} \qquad (25)$$

An example of use of the BART methodology was completed by using 10 EBR-II sensor signals. The BART system was trained using a training data set containing 1440 observation vectors. Out of the 1440 observation vectors, 129 of these were chosen to be used to construct a system model. The 129 vectors were also used to determine the height, h, of the angle domain boundary as well as the location of the BART reference point R for each of the sensors used in the experiment. To test the accuracy of the model 900 minutes of one minute data observation vectors under normal operating conditions were run through the BART system. The results of the BART system modeling accuracy are shown in FIGS. 12-16 and FIGS. 17-21 (BART modeled). The Mean Squared Errors for each of the sensor signals is shown in Table III.

TABLE III

BART System Modeling Estimation Mean Squared Errors for EBR-II Sensor Signals

Sensor Channel	Sensor Description	MSE of Estimation Error	Normalized MSE (MSE/H_3)	Normalized MSE (MSE/o_3)
1.	Primary Pump #1 Power (KW)	0.0000190	0.0000002	0.0002957
2.	Primary Pump #2 Power (KW)	0.0000538	0.0000004	0.0004265
3.	Primary Pump #1 Speed (RPM)	0.0000468	0.0000001	0.0005727
4.	Primary Pump #2 Speed (RPM)	0.0000452	0.0000001	0 0004571
5.	Reactor Outlet Flowrate (GPM)	8.6831039	0.0009670	0.1352974
6.	Primary Pump #2 Flowrate (GPM)	0.0571358	0.0000127	0.01.63304
7.	Subassembly Outlet Temperature 1A1 (F)	0.0029000	0.0000034	0.0062368
8.	Subassembly Outlet Temperature 2B1 (F)	0.0023966	0.0000027	0.0052941
9.	Subassembly Outlet Temperature 4E1 (F)	0.0025957	0.0000029	0.0050805
10.	Subassembly Outlet Temperature 4F1 (F)	0.0024624	0.0000028	0.00 1358